DOCUMENT RESUME

ED 217 403 CS 006 726

AUTHOR Kleiman, Glenn M.
TITLE Comparing Good and Poor Readers: A Critique of the Research. Technical Report. No. 246.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.; Toronto Univ. (Ontario).
PUB DATE Jun 82
CONTRACT 400-76-0116
NOTE 56 P.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Aptitude; *Comparative Analysis; *Listening Skills; Literature Reviews; *Reading Ability; Reading Comprehension; *Reading Difficulties; *Reading Research; Reading Skills; Research Methodology; Research Needs; *Research Problems; Suprasegmentals; Testing Problems

ABSTRACT

The many studies comparing good and poor readers have yielded few conclusive findings as a result of a number of problems. The first set of problems has to do with the choice of tests or tasks. There are many differences in the cognitive demands of reading and listening tasks children encounter in school. In listening, prosodic cues facilitate sentence analysis, and teachers provide external aids in retrieving relevant prior knowledge, focusing attention on main ideas, and monitoring comprehension. Determining main ideas and monitoring attention are especially important in taking advantage of the permanence of written language. As psychologists and educators turn their attention to reading comprehension, rather than individual word recognition and decoding, the differences between listening and reading warrant careful study. It is likely that causes of many reading problems will be found in the skills necessary in reading that differ from the skills children have mastered in listening. Another set of problems has to do with subject sampling and experimental design and measurement. Good discussions of these problems have been available for many years, but inadequate procedures continue to appear in published studies, and results from these studies continue to be accepted in review articles. Even when the procedures are adequate, there are serious problems in interpreting the results of studies comparing good and poor readers. The studies do not identify causal factors of reading disability nor do they recognize the great diversity of patterns of reading disability within each group of students. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED217403

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CENTER FOR THE STUDY OF READING

Technical Report No. 246

COMPARING GOOD AND POOR READERS: A CRITIQUE OF THE RESEARCH

Glenn M. Kleiman
University of Toronto

June 1982

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

The preparation of this report was supported by NIE under Contract No. HEW-NIE-C-400-76-0116 and by the University of Toronto. The author is now at Teaching Tools: Microcomputer Services, Palo Alto, California. To appear in K. Nelson (Ed.), Child Language (Vol. 5). New York: Gardner Press, in press.

CS206726

EDITORIAL BOARD

William Nagy and Stephen Wilhite
Co-Editors

Harry Blanchard
Charlotte Blomeyer
Nancy Bryant
Larry Colker
Avon Crismore
Meg Gallagher
Anne Hay
Asghar Iran-Nejad
Jill LaZansky
Terry Turner
Janet Williams
Paul Wilson

Michael Nivens, Editorial Assistant

Comparing Good and Poor Readers: A Critique of the Research

Many studies of children's reading have compared reading ability groups on measures of cognitive performance. The primary aim of this work has been to identify the underlying causes of children's reading problems. A large variety of measures have been used, including tests of perceptual discrimination, visual scanning, within-modality and between-modality matching, vocabulary knowledge, decoding, whole word recognition, short-term memory, memory for sentences, deductive and inductive reasoning, verbal and nonverbal IQ, and many more. The population of main interest has been children who have reading problems not attributable to neurological, physiological, emotional, general cognitive, or environmental factors. These children are often said to be dyslexic or to have specific reading disabilities. Since the definitions of these terms are subject to debate (Rutter, 1978; Rutter & Yule, 1975), labels such as below average, disabled, poor, problem, or retarded readers are used in many studies. The comparison children who do not have reading problems are typically called normal, good, superior, or skilled readers.

Studies comparing good and poor readers can be divided into three general categories according to the dependent measures used. One category consists of studies using measures of reading performance, such as number of comprehension questions correctly answered, number of errors in oral reading, or speed of reading. Studies using these measures fall into two subtypes. One subtype involves manipulations of aspects of the text, such as vocabulary difficulty, syntactic complexity, presence of illustrations, or use of adjunct questions. The other subtype involves comparisons of

reading performance before and after a training program. Some training studies look at poor readers only, but many include comparisons of reading ability groups.

The second category consists of studies measuring performance on subprocesses of reading, such as letter discrimination, word recognition, knowledge of letter-sound correspondences, or sentence parsing. Most of the studies in this category have focused on processes of word recognition. Many of these studies have simply tested for absolute differences between good and poor readers. Such differences are generally found; poor readers score lower than good readers on most measures of cognitive performance. The more interesting studies in this category have been concerned with interactions between reading ability and two or more measures of performance. That is, they look for patterns of differences--not only are poor readers' scores typically lower than good readers' scores on measure X, but the difference is greater than on measure Y.

The third category consists of studies measuring performance on cognitive processes that are not specific to reading. These studies are based on the view that poor readers' problems are not reading specific, but stem from a more basic or general cognitive deficit. Examples in this category are studies comparing good and poor readers on tests of visual discrimination, short-term memory span, or IQ. As in the previous category, there are studies that test for absolute differences between good and poor readers, and studies that test for interactions or patterns of differences.

Reviews of subsets of the extensive literature comparing good and poor readers can be found in many places, including a recent book by Vellutino (1979), volumes edited by Benton and Pearl (1978), Knights and

Bakkar (1976), and Waller and MacKinnon (in press), and Reading Research Quarterly articles by Golinkoff (1975-76), Samuels (1973-74), and Torgeson (1978-79). I will not re-review this literature here. Rather, I will focus on some problems that severely limit the interpretability and generalizability of much of this research.

In any of the types of studies described above, the researcher must: (a) decide on what tasks or tests good and poor readers should be compared; (b) obtain samples of good and poor readers and measure their performance; and (c) interpret the results. There are problems that arise, and have often been neglected or answered simplistically, at each of these steps. These problems are the focus of this chapter. The problems associated with each of these steps will be considered in the three main sections of this chapter.

Problems in Choosing the Tasks or Tests

A researcher's choice of the tasks to use in a study comparing good and poor readers rests on the assumed answers to two questions: (a) What types of knowledge and cognitive processes are required for skilled reading? (b) On which type of knowledge or processes are poor readers likely to be most deficient? That is, the choice is based on the researcher's views of skilled reading and reading disability.

A View of Skilled Reading

The view of skilled reading adapted in this chapter is well characterized by the following three quotes, all over 70 years old:

> Understanding a paragraph is like solving a problem. . . . The mind is assailed as it were by every word in the paragraph. It must select, repress, soften, emphasize, correlate and organize, all under the influence of the right mental set or purpose or demand. (E. Thorndike, 1917)

> To completely analyze what we do when we read would almost be the acme of a psychologist's achievements, for it would be to describe very many of the most intricate workings of the human mind. (Huey, 1908/1968)

> A reader or listener has at each moment but a limited amount of mental power available. To recognize and interpret the symbols presented to him requires part of this power; to arrange and combine the images suggested by them requires a further part; and only that part which remains can be used for framing the thought expressed. Hence the more time and attention it takes to receive and understand each sentence, the less time and attention can be given to the contained idea; and the less vividly will that idea be conceived. (Spencer, 1852/1881)

The three main characteristics of skilled reading captured by these quotes are: (a) It is a goal directed, flexible, cognitive skill. (b) It is complex, requiring the coordination of multiple processes and the use of knowledge of the language and the world in general. (c) Reading, like all cognitive skills, is constrained by limits of the human information processing system, such as short-term memory span and attentional capacity. These points are elaborated below.

Reading is goal directed and flexible in that good readers can use written texts in many ways. They can skim for main points or scan for particular information. They can read quickly or slowly, carefully or cursorily, silently or aloud. They can read for gist or for detail, to proofread or to memorize. They can read many types of materials, from

comic books to technical journals, from recipes to novels, from students' essays to Shakespeare's plays.

The many intricate workings of the human mind required for reading include general cognitive processes, such as perceptual discrimination, short-term memory storage and retrieval, serial order encoding, attention allocation and direction, long-term memory encoding and retrieval, and inferential processing. Also included are language comprehension processes, such as retrieving word meanings, parsing sentences, integrating word meanings, determining anaphoric references, and analyzing discourse structures. Reading-specific processes, such as using letter-sound correspondences and recognizing the visual forms of words, are also necessary. These processes interact in many ways. For example, the retrieval of appropriate world knowledge can facilitate meaning retrieval and sentence parsing (see Rumelhart, 1977).

According to this view, reading comprehension is not a passive accumulation of word or sentence meanings, nor is it simply an active attempt to reconstruct the meaning intended by the author. Rather, the reader has certain goals, either implicit or explicit, specific or general, and these goals influence the reading process. For example, when rapidly skimming text, readers rely on prior knowledge of the topic and a sample of the words in the text. They do not carefully process each word nor carefully parse each sentence. That is, they depend heavily on top-down or knowledge-driven processes. Alternatively, when reading carefully, readers will process each word and sentence will be more dependent on bottom-up or text-driven processing (Rumelhart, 1977).

This flexibility requires the inclusion of executive or control processes in an analysis of reading (Brown, 1980). These serve to determine the overall goal of reading, divide it into manageable or local subgoals, choose and integrate the processes to be used, and continually monitor their success, making adjustments in processing when necessary.

The human information processing system is limited in a number of ways (Newell & Simon, 1972), and the many subprocesses of reading must be coordinated to operate within these limits. For example, there are limits on how much can be perceived in a single fixation, how quickly the eyes can move, how many chunks of information can be held in short-term memory, and how quickly information can be retrieved from long-term memory. There is also a general limit in attentional capacity or cognitive resources. As Spencer pointed out over 120 years ago, if we attend to the individual pieces, we cannot attend to the meaning of the whole. In order to read well, the lower-level processes, such as word recognition, must function automatically (i.e., without requiring attention), so that attention can be directed to higher-order meanings (LaBerge & Samuels, 1974).

Where Might Poor Readers' Problems Lie?

Given this view of reading, the next question is: On what aspects of reading might poor readers be most deficient? There are many possibilities. They might be deficient in one or more of the general cognitive processes, such as perceptual discrimination, short-term memory storage, or long-term memory access. Or they might be deficient in reading or language-specific processes, such as word recognition, meaning retrieval or sentence parsing. Or perhaps poor readers lack the requisite knowledge

--they may not know the words or the concepts found in the texts they are expected to read. Or they may be adequate in all the necessary knowledge and individual processes, but deficient in applying and integrating them--that is, in the executive processes. Poor readers may, for example, overrely on knowledge-driven or text-driven processes (Spiro, 1979).

A large proportion of the studies comparing good and poor readers have focused on word recognition and decoding (for reviews see Barron, in press; Golinkoff, 1975-76; Vellutino, 1979).[1] This emphasis on word-level processing is based on two assumptions. One assumption is that, except for word recognition, reading and listening comprehension require identical cognitive processing (see Danks, 1974). The second assumption is that most children are fairly competent at listening comprehension by the time reading instruction begins. These assumptions lead to the view that reading problems most frequently stem from difficulty in recognizing written word or decoding written words to their spoken equivalents.

The assumption that listening and reading comprehension require identical cognitive processing (once word recognition is accomplished) is challenged below. It will be argued that there are important differences in the cognitive demands of naturalistic reading and listening tasks, and that reading is more demanding than listening on several higher-order cognitive processes that are not well developed in many children. These processes warrant consideration in attempts to identify the causes of reading problems.

Differences Between Written and Spoken Language

Consideration of the differences between written and spoken language has a long and respectable history, although these differences have generally been neglected in recent psychological, educational, and linguistic research. Plato, in the dialogue Phraedrus, pointed out that speakers can modify their communications to fit individual listeners, and listeners can influence the communication, but these options are not available to authors and readers. Aristotle, in the Art of Rhetoric (Book III, Chapter XII), pointed out that writing and speech differ in both function and style. His discussion included the greater precision and detail typically found in writing, the greater amount of repetition found in speech, and the availability of intonation in speech but not writing.

The Russian psychologist Vygotsky (1962, originally published in 1933) described many of the differences between writing and speech. He considered differences in sentence structure, precision, and detail ("In writing . . . we are obliged to use many more words, and to use them more exactly"). He discussed the effects of prosody and gestures on spoken communication, citing a passage from Dostoyevsky in which the same spoken word is said to be used with six different meanings. Vygotsky's description of the uses of the two modes of language is especially worth considering:

> Writing is addressed to an absent or an imaginary person or to no one in particular--a situation new and strange to the child . . . In conversation, every sentence is prompted by a motive. Desire or need lead to request, question to answer, bewilderment to explanation. The changing motives of the interlocutors determine at every moment the turn oral speech will take. It does not have

> to be consciously directed--the dynamic situation takes care of that. The motives for writing are more abstract, more intellectualized, further removed from immediate needs. In written language, we are obliged to create the situation, to represent it to ourselves. This demands detachment from the actual situation. (p. 99)

More recently, Olson (1976, 1977) and Rubin (1980) have discussed differences between the spoken language preschool children have mastered and the language children encounter in school. Olson emphasizes a distinction (similar to Vygotsky's) between interpersonal language and ideational language. Interpersonal language, the language of conversation, is familiar to young children. It is the language of action, used for requesting, questioning, and responding. It is closely tied to the immediate situation and to the desires and interests of the communicants. Ideational language, on the other hand, is not very familiar to children before they begin reading. It functions to communicate ideas, to describe and explain. It tends to be much more abstract, and to have less immediate relevance than interpersonal language.

Rubin (1980) presents a taxonomy of language experiences and a set of dimensions on which these experiences differ. The language experiences include engaging in conversation, listening to a radio, watching a play, reading a letter, reading a story, and a variety of others. The dimensions on which these differ are divided into two types, medium related and message related. Medium related dimensions include whether the message is written or spoken, whether the communicants can interact, whether they share a spatial and temporal context, whether the receiver is directly involved in the communication, and whether the referents of the message

are physically present. The medium dimensions are structure (vocabulary, syntax, and discourse organization), topic, and function.

One main thrust of both Olson's and Rubin's papers is that the cognitive processes mastered by preschool children in their language experiences may be qualitatively different from those required to understand much of the language encountered in school. In this section, I will limit consideration to "school language," and more specifically to language that presents information to be learned. I will argue that even when the goal of listening and reading is restricted in this way, important differences remain in the cognitive demands of understanding the two modes of language. First, I will consider differences due to the availability of prosodic information (intonation, rhythm, and stress or accent) in speech but not in writing. Then I will discuss some of the advantages teachers have over textbooks in making their presentations of information easily understood by children.

Prosody. Prosodic information is conveyed by patterns of pitch, loudness, and duration, and therefore cannot be directly represented in writing. As the linguist Bolinger has stated:

> The convergence of writing and speech virtually stops at the level of morphemes . . . Writing never really got around to providing a regular way of marking accent . . . Punctuation and capitalization serve as a rough guide to some of the rhythmic and intonation contrasts in speech, but much is left out. (1975, pp. 471-472)

Linguistic and acoustic analysis have shown that certain prosodic patterns tend to co-occur with certain aspects of speech. Prosody provides information that may be useful to listeners in determining:

(a) changes in the topic of discourse (Bolinger, 1975); (b) the ends of sentence, and whether they are statements, questions, or commands (Bolinger, 1975; Lefevre, 1970); (c) whether sentences convey direct or indirect speech acts (Sag & Liberman, 1975); (d) the ends of clauses within sentences, and, in many cases, phrases within clauses (Cooper & Sorenson, 1977; Grosjean, Grosjean, & Lane, 1979; Kleiman, Winograd, & Humphrey, 1979; Scholes, 1971; Sorenson, Cooper, & Paccia, 1978); (e) the words of a speaker wants to make prominent because they convey new or contrastive information (Bolinger, 1972; Chafe, 1974, 1976; Hornby, 1972; Lieberman, 1963); and (f) the referents of some pronouns (Maratsos, 1973).

Written language can be understood without prosody because prosodic information is usually redundant with syntactic, lexical, or semantic information, or is replaced by punctuation. Although language without prosody can be understood, the redundancy prosody provides may facilitate comprehension, and the lack of prosody in written language may contribute to reading problems (Kleiman et al., 1979; Read, Schreiber, & Walia, in press). The possibility that children have difficulty compensating for the lack of prosody in written text is supported by evidence that children tend to rely more than adults on prosodic cues, and are less able to use syntactic, lexical, and semantic cues (Hornby, 1971; Hornby & Hass, 1970; Read et al., in press). The strongest evidence for a role of prosody in understanding speech is in the use of prosodic cues for sentence parsing, and in determining the words the speaker intends to mark as focal. Some of this evidence is reviewed below.

Parsing sentences into meaningful phrases and clauses is an essential step in language comprehension. According to current models (e.g.,

Clark & Clark, 1977; Kleiman, 1975), language comprehension involves a limited capacity working memory that holds surface representations of input words. Various processes operate upon the words in working memory to parse them into constituents (phrases and clauses) and to determine the meanings expressed. In the comprehension of both written and spoken language, lexical, syntactic, and semantic information can be useful in parsing sentences (see Clark & Clark, 1977, chapter 2). Speech also contains useful prosodic information.

Once the meaning of a constituent has been determined, the individual words no longer need to be held in working memory, thereby freeing some of its capacity for new input. If one fails to parse sentences appropriately, comprehension will be impaired. In fact, there is evidence that parsing difficulty is often an aspect of reading comprehension failure. A common reading problem is that of reading "word by word" rather than chunking the words into meaningful phrases and clauses (Clay & Imlach, 1971; Golinkoff, 1975-76). This may be due in part to these children having difficulties compensating for the lack of prosodic cues in parsing written sentences (Kleiman et al., 1979; Read et al., in press).

Several types of prosodic cues to phrase and clause boundaries have been identified in acoustical studies. Cooper and Sorenson (1977) found evidence that these boundaries tend to be marked by a specific pattern of pitch change. Klatt (1976) and Sorenson et al. (1978) found that phrases are marked by an increase in the duration of their final syllables. Scholes (1971) argued that the relative peaks in loudness provide the most reliable cues to syntactic boundaries. In addition, pauses in speech provide information of potential use in parsing (Grosjean, Grosjean, & Lane, 1979).

These studies suggest that pitch, duration, loudness, and pauses can all provide useful information. Which dimension dominates may vary according to the speaker and the structure of the sentence.

The usefulness of prosodic cues in parsing is most apparent in cases of surface structure ambiguity. For example, the sentence I fed her dog biscuits has two possible readings: either she was fed dog biscuits or her dog was fed biscuits. In speech, these two readings would be reflected in different prosodic patterns which would enable the listener to determine whether the appropriate parsing is I fed her / dog biscuits or I fed / her dog / biscuits. Lehiste (1973) provides evidence that listeners can use prosodic information to determine the intended meanings of such sentences.

Prosodic cues to sentence structure are also available in sentences that are not ambiguous. In a study by Scholes (1971), pairs of words were used that, when placed in different contexts, either were within the same clause or had a clause boundary between them. For example, the word pair spotted plants appeared in the following two sentences:

If you find your flowers spotted plant them in the sun.

If you find your spotted plant let me know.

Tape recordings were made of speakers reading each of the sentences aloud. The word pairs were then excised from the sentences and played to subjects who were asked to judge which sentence each word pair was in when it had been recorded. Subjects did significantly better than chance.

Several other studies have provided evidence for the use of prosody in sentence parsing by creating sentences in which there is a mismatch between prosodic and syntactic information. This is done by using sentence pairs that have a string of words in common, but different constituent

boundaries within the string, as in the spotted plant example given above. Both sentences are recorded with normal intonation, and then the common word string is spliced from one context to the other. These studies have yielded three main findings. One is that subjects' recall errors generally consisted of changes in wording such that the syntactic structure of the reported sentence fit the prosodic pattern that was actually presented (Darwin, 1975; Wingfield, 1975; Wingfield & Klein, 1971). That is, subjects resolved the discrepancies between intonation and syntax by altering the syntactic structures of the sentences. This may be related to Garnes and Bond's (1975) finding that misperceptions of natural speech occur on phonemes, syllables, words, and phrases, but that stress and intonation patterns are rarely misperceived.

The second result is based on the finding that with normal spoken sentences, subjects tend to report accurately the location of interrupting stimuli (such as clicks) when they occur at syntactic boundaries, but tend to report them inaccurately when they occur within syntactic units (Fodor & Bever, 1965). In sentences in which syntax and prosody mismatched, interrupting stimuli that occurred at the boundary marked by prosody were reported most accurately (Wingfield & Klein, 1971; see also Geers, 1978).

The third relevant finding on the effects of misleading prosody has been recently reported by Read et al. (in press). They trained 7-year-old children to listen to sentences and then repeat the subject noun phrase only. When presented with normal sentences, in which prosody and syntax match, the children correctly repeated the subject noun phrase 83% of the time. When prosody and syntax mismatched, the children were correct on only 30% of the sentences. Moreover, 78% of the errors in children's

repetitions ended at the misleading prosodic cue. In the same task, adults were much less affected by misleading prosody.

Prosody also provides cues to the topic of the sentence (Hornby, 1971), the new information (Clark & Haviland, 1977), and information the speaker believes contrasts with the listener's expectations or prior information (Chafe, 1976). These cues take the form of sentence stress or accent. In an acoustical analysis, Lieberman (1960) found that fundamental frequency, relative amplitude, and duration are all correlates of stress. Lieberman (1963) demonstrated that speakers produce the more informative (i.e., less predictable) words with more stress than other words.

A clear case of sentence stress is found in question answering. Consider the following sentences spoken with the capitalized word stressed:

(1a) JOHN stole the picture.

(b) John STOLE the picture.

(c) John stole the PICTURE.

In each case the stressed word would be the one carrying the new information while the rest of the sentence specifies the given information. That is, sentences in (1) could be answers to questions in (2).

(2a) Who stole the picture?

(b) What did John do with the picture?

(c) What did John steal?

Hornby (1971) studied children's use of stress and syntactic cues in determining the topic of sentences. He presented children with active, passive, cleft, and pseudocleft sentences, and sentences with contrastive stress. The children's task was to select a picture that shows the action described in the sentence. The sentence and pictures were designed so that

the choices would reflect the children's views of the topic of the sentence. The results led Hornby to conclude that stress is the primary device children use in determining the topic of sentences (see also Hornby & Hass, 1970).

Teachers versus textbooks. In addition to the use of prosody, two other aspects of teachers' presentations make them very different from textbook presentations: Teachers and students interact during the presentation, and teachers have some knowledge of what the students already know and do not know and how easily they can understand new material. Both of these aspects of speech were contrasted with writing by Plato:

> Written words seem to talk to you as though they were intelligent, but if you ask them anything about what they say, from a desire to be instructed, they go on telling you just the same thing forever . . .
>
> Speech can be varied so that it is appropriate to each nature . . . addressing a variegated soul in a variegated style . . . and a simple soul in a simple style.

The French novelist Sartre (1964) provides an analysis from a very different perspective, but points out the same distinctions between natural spoken language and the language of books. Recalling his shock the first time his mother read him a story, he writes:

> I was bewildered: who was telling what and to whom? My mother had gone off: . . . I didn't recognize her speech . . . A moment later, I realized: it was the book that was speaking. Frightening sentences emerged from it: they were real centipedes, they swarmed with syllables and letters . . . Rich in unknown words, they were enchanted with themselves and their meanderings without bothering about me. Sometimes they disappeared before I was able to

> understand them; at other times I understood in advance; and they continued to roll nobly to their end without sparing me a single comma. That discourse was certainly not meant for me. (p. 46)

It is well documented that speakers modify their language to suit their listeners (e.g., Snow & Fergusen, 1977) and that in interactive situations listeners provide a great deal of feedback to speakers (Wilkinson, 1971). In classrooms, a very common "teaching cycle" consists of the teacher asking a question, one or more students responding, and the teacher then evaluating or modifying the response (Bellack, Davitz, Kliebard, & Hyman, 1963; Sinclair & Coulthard, 1975).

In a study of differences between listening-to-learn and reading-to-learn, Schallert and Kleiman (1979) obtained samples of expository texts written for children, and tape-recordings and observations of teachers presenting comparable material to their classes. The children's reading materials had been adapted from materials originally intended for adults. The teachers used the adult materials as a basis for preparing their presentations.

These language samples have been used to identify some of the ways teachers use their knowledge of the children and the flexibility of oral presentations to make their lessons easier to understand than comparable material presented in textbooks. Teachers can adjust the amount and complexity of the material covered, and the vocabulary and sentence structures used. In addition, teachers can provide external aids to help children with three processes, each critical to comprehension and learning. Schallert and Kleiman (1979) refer to these processes as activating relevant prior knowledge, focusing attention on main ideas, and monitoring comprehension.

Children with reading difficulties may lack adequate background knowledge (Anderson, 1977), or fail to make use of the knowledge they do have (Spiro, 1979). When teachers orally present lessons, they often guide the students in retrieving and using relevant prior knowledge. Teachers can check the students' prior knowledge to fill in missing information, correct erroneous information, and relate new information to things the students already know. For example, in Schallert and Kleiman's language samples, one of the teachers began her presentation of a lesson on Sequoia trees as follows:

Teacher: Today we are going to learn about something that's the oldest and the biggest living thing that we know of. The oldest and the biggest. Now think just a minute before you get your hand up. The oldest and the biggest. What do you think it is--Jeff?

Student: Dinosaur.

T: Why is dinosaur not a good answer?

S: Not living.

This type of interchange continues with students suggesting elephant, whale, shark, and the earth, until:

S: Trees. Trees are living.

T: All right. Say it again. Listen again. Heidi's got the answer over here. Say it again.

S: Sequoia tree.

T: Sequoia trees. How many of you've ever heard of a sequoia tree?

With this brief introduction to the lesson, the teacher has done several things which may help her students understand and learn the material. First, she began by finding out about the children's prior knowledge. This provided an opportunity to correct their initial responses, and, in so doing, to make clear the characteristics that are central to the discussion. Moreover, it enabled her to remind the children of information they already knew, and to contrast the new information with the already known. When one student gave the correct answer, the teacher directed the class' attention to that child, had the child repeat it, and then repeated the answer herself. The teacher then went on to find out more about the students' prior knowledge by asking how many have heard of sequoia trees.

Another aspect of skilled reading that is difficult for many children is determining the main ideas of a passage. In studies with subjects from third grade through college, Brown and Smiley (1977a, 1977b) found large developmental differences in the ability to determine the importance of structural units of prose passages. When given time to study the passage, subjects who were able to determine which parts were important focused their attention on those parts, while subjects who were unable to pick out the main ideas distributed their study time over all the information in the passage. Eamon (1978-79) also presents evidence that distinguishing main ideas from peripheral information can be difficult for poor readers. In presenting information, teachers explicitly point out main ideas, and they provide cues to importance by intonation patterns, amount of repetition, and phrasing. All of this may make the task of determining main ideas much easier when listening to teachers' presentations than when reading.

Effective monitoring of one's own comprehension, determining whether or not one has understood the information, is another critical process in language comprehension. Recent studies have shown that children are often poor at comprehension monitoring. For example, in a study by Markman (1979) elementary school children were asked to act as consultants to help evaluate essays. Each essay had a blatant contradiction. For example, part one of one essay read: "Fish must have light in order to see. There is absolutely no light at the bottom of the ocean. It is pitch black down there. When it is that dark, fish cannot see anything. They cannot even see colors. Some fish that live at the bottom of the ocean can see the color of their food." Children often judged these essays as making sense and being easy to understand. Further evidence is provided by studies of referential communication in which children serving as listeners were instructed to ask questions if they needed more information (Cosgrove & Patterson, 1977; Ironsmith & Whitehurst, 1978; Patterson, Massad, & Cosgrove, 1978). The children, particularly those below the fourth grade, often failed to request further information, even when the original message was completely uninformative.

Teachers, when presenting information, frequently monitor the students' comprehension by asking questions. They also note looks of puzzlement and drifts of attention. When the students are not adequately comprehending, the teacher will repeat and rephrase information or fill in necessary background information. This external monitoring, repetition, and further information makes the need for children to monitor their own comprehension much less important in listening-to-learn than in reading-to-learn.

Determining main ideas and monitoring comprehension are relevant to one advantage that writing has over speech. Writing is permanent, and therefore readers control how they sample information from the text. Readers can preview the material, choose to read some parts slowly and carefully and to skim others, and reread at will.

There is evidence that good readers take advantage of the options afforded by the permanence of written texts. Tinker (1958) reported that the rate of reading decreases as the text becomes more difficult. Furthermore, the pace is not simply set and then maintained throughout the text. Skilled readers slow down for important or confusing passages and speed up for easy or unimportant ones. Taylor (1957) reported that 15% of all eye movements in college-level readers are regressive. The use of this rereading option may be crucial for skilled reading. Skilled readers proceed rapidly, hypothesizing about what will come next and integrating what is read with previous parts of the text. The rereading option enables them to do this without taking too large a risk of misinterpreting or failing to comprehend, since they can go back and reread when necessary. Wanat (1971) demonstrated that regressive eye movements are likely to occur when the text does not match readers' expectations. He compared adults' eye movements while they read two types of sentences, agentive passives (e.g., The ball was hit by the boy) and locative passives (e.g., The ball was hit by the park). Since passive sentences usually specify the agent at the end, readers are more likely to expect an agent, such as boy, than a location, such as park. Wanat found more regressions and longer regression durations with the locative passives than the agentive passives. Also, the regressions usually occurred after the locative and were directed

back to the word by. Further evidence that the nature of the text influences eye fixation patterns is provided by Rayner and McConkie (1976), Carpenter and Just (1978), and Just and Carpenter (1980).

The ability to sample the text efficiently is an important reading skill, one that differs from any skills used in listening. A study by Neville and Pugh (1976-77) provides evidence that some readers in the middle grades do not make good use of the potential to sample information. They tested fifth graders on three types of cloze tests: a regular reading test, a restricted reading test, and a listening test. On the listening and restricted reading tests, information about the words following the missing one was not available. On the regular reading cloze test, this information was available. However, only the better readers seemed to make use of it. The poor readers' performance was equivalent on all three types of tests, and their errors on the regular reading test were consistent with the preceding context but were sometimes inconsistent with the words that followed. The good readers' performance on the regular reading test was superior to the other two tests, and their errors were consistent with both the preceding and following context.

In order to efficiently sample information from written texts, readers must continuously evaluate what they are reading to determine if it is important and needs to be read carefully, and they must constantly monitor their own comprehension to determine if they are understanding the text sufficiently. These evaluating and monitoring processes are often difficult for children. The demands for them are much greater in reading than in listening.

To summarize the main argument presented in this section, there are many differences in the cognitive demands of reading and listening tasks children encounter in school. In listening, prosodic cues facilitate sentence analysis, and teachers provide external aids in retrieving relevant prior knowledge, focusing attention on main ideas, and monitoring comprehension. Determining main ideas and monitoring attention are especially important in taking advantage of the permanence of written language. There is evidence that each of these aspects of reading present difficulties for children. As psychologists and educators turn their attention to reading comprehension, rather than individual word recognition and decoding, the differences between listening and reading warrant careful study. It is likely that the causes of many reading problems will be found in the skills necessary in reading that differ from the skills children have mastered in listening.

Problems in Subject Sampling and Measurement

There are serious sampling and measurement problems in many of the techniques used in studies comparing good and poor readers. These problems have long been known. Excellent discussions of them can be found in Campbell and Stanley (1963) and R. Thorndike (1963). However, these methodological faults continue to appear frequently in published studies. In this section, some of the most frequent and critical problems will be discussed briefly. The reader interested in more extensive and technical discussions is referred to Applebee (1971), Campbell and Stanley (1963), Chapman and Chapman (1973, 1974), Cronback and Snow (1977), and R. Thorndike (1963).

The most frequent and critical problems in studies comparing reading ability groups include: (a) lack of comparability across studies and

restricted generalizability of findings; (b) misguided procedures for matching good and poor readers on IQ and other variables; (c) low reliability of measures and neglect of statistical regression; and (d) neglect of effects attributable to the scale of measurement used. Each of these is discussed further below.

Lack of Comparability and Generalizability of Findings

Samples of poor readers in different studies are often not comparable because they were obtained from different populations, or different selection criteria were used. Poor readers selected from regular classes may not be as severely disabled as those selected from remedial reading classes. Both of these groups may differ from poor readers referred to clinics, and different types of clinics (e.g., psychological versus neurological) tend to receive children with different types of problems.

In most studies, samples of poor readers are obtained from schools, and poor readers are identified on the basis of standardized reading test scores. Unfortunately, the fact that different reading tests emphasize different types of materials and questions is usually neglected. For example, some tests heavily weight individual word reading, others weight literal comprehension questions, and others weight inferential questions (i.e., questions requiring information beyond that stated in the text). Some tests require a great deal of oral reading, others require no oral reading at all (see Farr, 1969; MacGinitie, 1973). Jerrolds, Calloway, and Gwaltney (1971) showed that different tests will yield different numbers of children classified as disabled readers. They also found that the groups of children identified as disabled according to different reading tests showed different patterns of performance on verbal and performance IQ

tests. That is, different reading tests did not just alter the number of children classified as poor readers; the poor readers' patterns of performance on other cognitive tests were also altered.

Differences among tests are particularly critical when cross-age comparisons are made. Even when tests from the same series are used, the skills measured at different grades are not the same. For example, the Stanford Diagnostic Reading Test has four levels, one each for Grades 1-2, 3-4, 5-8, and 9-12. The test for Grades 1-2 heavily weights measures of auditory discrimination and auditory vocabulary. The test for Grades 3-4 heavily weights phonetic and structural analyses of written words and word parts. The tests for the older children heavily weight literal comprehension, inferential comprehension, and rate of reading. Children classified as having reading problems on the basis of one level of a test should not be assumed to have problems with the same aspects of reading as children classified as having reading problems on the basis of a different test, even if it is another level from the same test series.

The criteria used to classify children as poor readers will also affect the nature of the sample obtained. The most common procedure is to classify children as poor readers if their reading test scores are two or more years below their age-appropriate grade placement. This criterion makes interpreting patterns across grades difficult. The percentage of children fitting this criterion increases with grade level, ranging from less than 2% at the beginning of third grade to nearly 30% by the end of ninth grade (Applebee, 1971). By most standards, a ninth-grade child reading at the seventh-grade level is not as severely disabled as a third-grade child reading at the first-grade level. To make samples of poor

readers more comparable across grades, criteria have been proposed that would classify as poor readers the same proportion of children at each grade (e.g., Jerrolds et al., 1971).

Problems in Matching Good and Poor Readers

Most researchers aim to identify differences between good and poor readers that are specific to reading, not attributable to IQ differences. These researchers have tried to identify children who are reading more poorly than would be expected on the basis of their IQ scores. Rutter and Yule (1973, 1975), in an epidemiological study, obtained strong evidence that reading problems not attributable to IQ occur in an educationally significant proportion of the population. They also found that the populations of children with reading problems attributable to low IQ and of children with reading problems not attributable to low IQ differ in many ways, such as male/female ratios, severity of spelling problems, and prognosis for improvement. Despite the demonstrated existence of the population, attempts to limit the sample of poor readers to those whose reading problems are not due to low IQ have resulted in some misguided procedures.

Two procedures have been widely used. One is to use a minimal IQ criterion, often 90. Only children whose reading scores are below a set criterion and whose IQ scores are above the minimal criterion are selected for the study. Alternatively, the subject selection criterion is set as a difference between IQ and reading scores. Typically, both scores are converted to chronological age (i.e., the age at which the obtained score would reflect average performance), and if the reading score is two or more years below the IQ score the child is classified as having a specific reading disability.

The problems with these procedures stem from the fact that IQ and reading scores are highly correlated. IQ tests that emphasize verbal skills overlap in content with reading tests. In fact, factor analyses of both verbal IQ tests and reading tests generally yield vocabulary knowledge as the most heavily weighted factor. Even nonverbal IQ tests show positive and significant correlations with reading tests. Yule, Rutter, Berger, & Thompson (1974) reported correlations between nonverbal IQ and reading scores for five populations. The five correlations, each based on scores from over 1100 subjects, are all greater than 0.6.

Since IQ and reading scores are positively correlated, children with low reading scores but not low IQ scores do not comprise a representative sample of poor readers. Therefore, the generality of findings from such a sample is restricted. More critically, when two measures are highly correlated, selecting a sample low on one measure but not low on the other results in selecting subjects with scores that have large errors of measurement (Calfee, 1976, pp. 25-34; Campbell & Stanley, 1963). In the case of children with low reading scores but not low IQ scores, retesting will result, on average, in higher reading scores and lower IQ scores than on the initial test. If, as is often done, the good or average readers are selected to match the poor readers on IQ, the match will not be as good as it appears; the poor readers' IQ scores will have large errors of measurement, and on retesting, the average IQ of these poor readers will be found to be lower. Similar problems occur when good and poor readers are matched on other variables, such as decoding ability or vocabulary knowledge (see Calfee, 1976, pp. 25-34).

Using a criterion based on differences between IQ and reading test scores is not a satisfactory solution to this problem. The basic problem with this procedure is that difference scores of correlated measures have low reliability. When difference scores are taken, the variability common to the two original measures (reflected in the correlation between the measures) cancels, while the errors in measurement (reflected for each measure by one minus the reliability coefficient) accumulate. The reliability of a difference score measure can be calculated by:

$$(3) \quad r_{diff} = \frac{\frac{1}{2}(r_{11} + r_{22}) - r_{12}}{1 - r_{12}}$$

where r_{diff} is the reliability of the difference score, r_{11} and r_{22} are the reliabilities of the two original measures, and r_{12} is the correlation of the two original measures (Thorndike & Hagen, 1977, Chapter 3). This formula shows that as the correlation between the two measures approaches their average reliability, the reliability of the difference score approaches zero. To give one example with values typical of those for reading and IQ scores, if the average reliability of the two tests is 0.8, and the correlation of the two tests is 0.6, then the reliability of the difference score measure is 0.5. This is not adequate reliability for most classification purposes. A large proportion of the children will be incorrectly assigned in a study that uses this procedure to assign subjects to reading ability groups.

The best procedure to use when IQ and reading ability need to be separated is to use an appropriate regression equation to predict children's reading scores on the basis of their IQ scores, and select those children whose reading scores are substantially below the predicted score (see

R. Thorndike, 1963; Yule, et al., 1974). This procedure should be used with IQ and reading tests that have minimal content overlap, and that are both highly reliable.

Reliability and Regression Problems

Poor readers, by definition, are at the low end of the distribution of reading scores. The extreme scores on any measure have the largest error of measurement. On any test with less than perfect reliability, regression to the mean will occur on retesting. Children with scores above the mean on the initial testing will score, on average, closer to the mean (although still above it) on retesting. Likewise, children with scores below the mean will tend to score closer to, but still below, the mean on retesting. The magnitude of the expected regression to the mean is determined by the reliability of the test (the higher the reliability, the smaller the regression effect), and the discrepancy of the initial score from the mean for the population (the larger the discrepancy, the larger the regression effect). More specifically, the best prediction of the change in score on retesting, assuming no other influences besides statistical regression, can be calculated by:

$$(4) \quad C = (M - S)(1 - r),$$

where C is the change from initial test to retest, M is the population mean, S is the score on the initial test, and r is the reliability of the test.

A widely cited study by Cromer (1970) provides examples of both inappropriate matching procedures and resulting problems due to statistical regression. Cromer's study is used by Gibson and Levin (1975) and Golinkoff (1975-76) as major support for the claim that there are poor readers whose problems are not in reading individual words, but are in sentence

organization. Most of the following discussion of Cromer's study is based on Calfee (1976).

Cromer's subjects were junior college students. He purported to divide the poor readers into two groups that were equally poor in reading comprehension, and to match each group with good reading comprehenders. The deficit poor readers also scored poorly on a vocabulary test, and hence word reading problems were assumed to be the cause of their comprehension problems. The difference poor readers scored as well on the vocabulary test as did their good reader matches, and hence it was assumed their comprehension problems stemmed from word-by-word reading or sentence organization difficulties.

To test the assumptions about deficit and difference types of poor readers, Cromer administered reading tasks to each group. Each task involved reading stories and answering multiple choice questions. The conditions differed in the ways the stories were presented. The two conditions relevant to this discussion were: (a) regular sentence presentation, and (b) phrase presentation, in which the sentences were divided into meaningful phrases.

Since the difference poor readers were thought to be poor at sentence organization, one of Cromer's predictions was that the pre-organized phrase presentation would facilitate reading comprehension for this group. Since the good readers and the deficit readers were thought to have adequate sentence organizational skills, their comprehension should be the same for the regular sentence and phrase presentation conditions. Cromer claims that the results of his study support these predictions.

Cromer's study is subject to severe problems in the subject selection procedures, much like those that arise when one attempts to match good and poor readers for IQ. The selection procedures used result in regression effects, which were not controlled for in the experimental design.

Reading comprehension and vocabulary measures are highly correlated. Selecting subjects who are high in one and low in the other (Cromer's difference poor readers) results in a high error of measurement. As a result of regression to the mean, on a retest the difference poor readers' vocabulary scores would decrease, while their comprehension scores would increase. They would no longer be well matched on vocabulary with their good reader controls, nor well matched on comprehension with the deficit poor readers. In fact, Cromer reports that in the regular sentence condition, 6 of the 16 difference poor readers had higher comprehension scores than their matched controls. Since this condition was basically a replication of the original measure used to assigned subjects to groups, it shows that the difference poor readers and their matched controls were not reliably assigned. Cromer ignored this problem and simply reassigned these subjects and reanalyzed the data. Further problems with this study are discussed by Calfee (1976).

The neglect of statistical regression is sufficient to make drawing of any conclusions from this study questionable. The same serious error can be found in other studies following Cromer's deficit and difference group distinction (e.g., Levin, 1973).

Neglecting statistical regression is a critical flaw in many studies using test-retest designs. Poor readers should be expected to improve their scores on retesting even without any training or other manipulations.

The more extreme the poor readers' initial scores, the larger the regression effect. The best solution to this problem is to divide the poor readers into treatment and control groups randomly. Both should show equal regression effects on retesting. Any differences between the two groups on retesting can then be attributed to the experimental treatment or training procedure (Campbell & Stanley, 1963; R. Thorndike, 1963).

Scale of Measurement Problems

Studies testing for absolute differences between good and poor readers, or for interactions between reading ability and performance on two or more measures, are also subject to scale of measurement problems. The simplest, most recognized, of these problems are ceiling and floor effects. If the task is very easy or very difficult, so that both good and poor readers score very high or very low, differences between the groups may be obscured. Ceiling and floor problems can also yield spurious interactions. For example, if poor readers perform better on test X than on test Y, while good readers perform at ceiling on both tests and therefore cannot do better on X than on Y, a statistical interaction may appear in the data analysis.

Statistical interactions are also of questionable interpretability when, as is typical in studies comparing good and poor readers, the groups are not equated on any baseline measure. One example of this problem is found in a set of studies on good and poor readers' use of sentence contexts in recognizing words. Some models of reading (e.g., Smith, 1973) claim that good readers make better use of contextual information than poor readers. In fact, in these models, use of contextual information is viewed as one of the main determinants of reading ability. Stanovich (1980) makes

the opposite claim that skilled readers make less use of prior sentence context in recognizing words while reading.

Stanovich's claim rests largely on his interpretation of three experimental studies by Perfetti, Goldman, and Hogaboam (1979); Schvaneveldt, Ackerman, and Semelar (1977); and West and Stanovich (1978). In all three studies, the time it took readers of various ability levels to perform a word recognition task was measured. The main comparison in each study was the time to recognize words with and without potentially helpful contexts. In each case, the absolute descrease in reaction time in the context condition, as compared to the no-context or control condition, was greater for the lower ability readers. However, there are problems in using this data to support the view that poor readers generally make greater use of context than good readers.

One major difficulty stems from the fact that in all cases the group with larger context effects also took longer to recognize the words both with and without context. For example, in the Schvaneveldt et al. (1977) study, the mean reaction times to recognize words in the relevant conditions were 1164 msec for the younger group, and 916 msec for the older group. The context effects (average time with context minus average time in the control condition) were 94 msec and 49 msec for the younger and older groups, respectively. This suggests that the apparently higher context effect may be a function of higher baseline time, which allows more possibility for any facilitory effect.

More technically stated, the problem stems from the fact that interactions may depend on the scale of measurement (Winer, 1971, pp. 449ff). Of the three research reports cited above, only Perfetti et al. (1979)

show any awareness of this problem. In the text of their paper, Perfetti et al. report a significant context by reading ability interaction, with the less able readers showing the larger context effect. However, in a footnote they report a second analysis in which they transformed the data to reduce the inequality of variances between reading ability groups. In this analysis, there was no context by reading ability interaction. More recent work provides evidence that, given comparable baselines, good and poor readers do not differ in their use of context to facilitate word recognition.

Chapman and Chapman (1973, 1974) provide more technical discussions of problems of developing measures that enable interpretation of interactions involving ability groups. In general, interactions of this sort are not directly interpretable unless the group that is superior on one relevant measure is inferior on another (i.e., a cross-over interaction), or the two groups are equal on a relevant baseline measure.[2] Since good readers are superior to poor readers on most cognitive measures, these conditions are rarely met in research comparing good and poor readers.

Problems in Interpreting Empirical Findings

In the previous section, I discussed some of the most common sampling and measurement problems found in studies comparing good and poor readers. In this section, the question to be considered is: Assuming appropriate methodology has been used, what conclusions can logically be drawn from findings of differences between good and poor readers? That is, once a difference between reading ability groups has been established, what are we to make of it?

The primary aim of the research considered in this chapter is to identify the underlying cognitive causes of reading disability. When poor readers are found to do less well than good readers on a cognitive task that can be related to reading, it is tempting to draw causal inferences. However, such inferences are likely to be erroneous.

A difference between good and poor readers might be a symptom of the actual underlying causal component, but not provide any useful information about it. A good example of a symptomatic difference is found in eye movement research. It has long been established that the eye movement patterns of good and poor readers differ (Tinker, 1958). However, attempts to train poor readers to move their eyes like good readers have not been successful in improving reading skill. Current models view eye movements as reflecting underlying cognitive processes (e.g., Just & Carpenter, 1980). From this current perspective, training eye movement patterns would not be expected to improve reading; it would be treating the symptom, rather than the cause.

Differences between good and poor readers may also be secondary effects of reading ability, rather than causes of reading ability. That is, deficits in poor readers may well be caused by their reading problems, rather than being causes of them. Or the reading problem and the other deficit(s) may cyclically reinforce one another. A good example here is vocabulary knowledge. Is lack of an adequate vocabulary a cause of reading problems? Or, since vocabulary is improved by reading, do reading problems cause poor vocabularies? Or is it some of each?

Finally, cognitive deficits may be correlated with reading problems without causal connections in either direction. This is very likely, since reading ability is related to IQ, socioeconomic status, and quality of schooling, and these factors can be expected to have wide-ranging effects.

Comparisons of reading ability groups do not provide a direct way of determining whether an obtained difference stems from a component of reading that causes reading problems, is a symptom of a causal component, is a secondary effect of reading ability, or is due to a process that is correlated with reading ability but not causally related to it. In some cases, causal differences can be separated from the other differences by data from training studies. If poor readers are deficient on cognitive process X (as compared to good readers), and this causes reading disability, then training poor readers on X should improve reading performance (given certain assumptions). If X is a symptomatic, secondary, or correlated difference, then training on X should not improve reading ability (see Fleisher, Jenkins, & Pany, 1979, and Weaver, 1979, for recent applications of this logic). Training studies have been fruitful in research on memory development and disabilities. For discussions of this work, and methodological suggestions, see Belmont and Butterfield (1977) and Brown and Campione (1979). Ryan (in press) has also advocated adapting the "instructional method" of studying memory to the study of reading.

The final problem to be discussed is perhaps the most critical, having to do with the basic assumptions about reading disability that underlie comparisons of reading ability groups. An excellent discussion of this problem can be found in Applebee (1971). He describes the general situation as follows:

> Any investigation of the problem of reading disability begins with a set of scores $\underline{X}_1$ to $\underline{X}_n$ on measures the investigator expects will be relevant to reading ability, and with a score $\underline{Y}$ which indicates performance on a reading criterion . . . Any analysis carried out on the $\underline{X}$'s in an attempt to predict $\underline{Y}$, or to describe the differences between groups specified in terms

> of relative scores on Y, is predicated, consciously or not, on some mathematical model of the functional relationship between the two sets of variables. (p. 99)

The problem is that a variety of models are logically possible, and the models assumed for group comparisons may be invalid.

According to the simplest model, one and the same cognitive component is responsible for all reading disabilities. This view underlies searches for a single cause of all reading problems (that are not attributable to low IQ, environmental, or neurological factors). Such searches have typically focused on a basic cognitive process such as visual perception, intersensory matching, serial order encoding, or short-term memory functioning (see Vellutino, 1979, for a review and critique of this work). There have also been proposals that the usual cause of reading problems are deficits in language-specific processes (e.g., speed of lexical access), reading-specific processes (e.g., decoding fluency), executive or control processes (e.g., setting the goal as word pronounciation and not attending to meanings), or lack of an adequate knowledge base.

If this simplest model held, group comparisons might be able to isolate the causal component; all poor readers would be deficient on it, and no good readers would be. More specifically, the distributions of good and poor readers' scores on a measure tapping this component would not overlap (except for measurement error). Scores on tests tapping other components should overlap, and on components that are not affected by the critical causal one, good and poor readers should not differ. Unfortunately, the available data do not take this form. Distributions of good and poor readers' scores generally have a great deal of overlap. The average score of poor readers is lower than the average score of good

readers on most measures of cognitive performance. These data, combined with the diversity of the knowledge and cognitive processes required for skilled reading, make it unlikely that there is a single cause of most reading problems.

According to a slightly more complex model, each reading problem is caused by a deficit in a single cognitive component, but the component varies across individuals. This view, coupled with an assumption that there are a small number of different causal components, underlies attempts to develop taxonomies of reading problems. If this model holds, group comparisons will show poor readers to score lower than good readers on a variety of measures. This might lead to the erroneous conclusion that poor readers are typically deficient on multiple components, when each individual is actually deficient on only one.

It is also possible that each reading problem is caused by a deficit on a single component, but that the deficit can be on any of the requisite processes or knowledge bases. If this is the case, it will be difficult to develop a usable taxonomy, and group comparisons will obscure the specific deficits of individuals.

As Applebee (1971) points out, the actual situation may be best characterized by a multiple regression model. According to this model, each of the many cognitive components of reading contribute to determining one's reading level. The various components may be differentially weighted, with certain ones contributing more of the variance in reading ability. In this view, reading disability can be caused by a slight deficit on many components, a moderate deficit on several components, or a large deficit on one or two components. That is, very different patterns of performance

across cognitive components would be found for different poor readers. Some might have a localizable deficiency, others a global one. Group comparisons would obscure such patterns.

In the final model to be considered, it is not simply the levels of abilities on cognitive components that matter, but patterns of abilities and how they are used by the control processes. For example, someone who is poor in individual word reading would usually be expected to be a poor reader. However, this may not be the case if the executive processes can use context and prior knowledge to compensate for the deficit in word reading. That is, good and poor readers might be equally capable on many components, but differ in certain key combinations, or in the way the control processes use strengths to compensate for weaknesses. Stated differently, there may be many different reading strategies, any of which can be successful for some individuals. Group comparisons will not provide information about the various possible strategies.

Given our current knowledge of the complexity of reading, it is likely that one of the more complex models (i.e., the multiple regression or patterns of abilities models) will be necessary to characterize reading disabilities. If one of these models holds, reading ability group comparisons will obscure rather than elucidate the patterns of reading disabilities. Perhaps, so much research effort has yielded so little progress because the assumed simple models are inappropriate.

The apparent solution is to focus on detailed analysis of individual reading problems. Information about many individual cases is necessary to determine which of the possible models of reading disability is most appropriate. This approach of going from the detailed study of individual

cases to general principles is the inverse of the nomethetic approach of group comparisons. The detailed study of individuals within a cognitive framework has been very productive in Newell and Simon's (1972) work on problem solving. It is also a central feature of the cognitive-functional approach (Meichenbaum, 1976; Ryan, in press).

Summary

The large number of studies comparing good and poor readers have yielded few conclusive findings. This is due to a variety of serious problems, many of which have been discussed in this paper. The first set of problems discussed, those having to do with the choice of tests or tasks, may be remedied by research within the next few years (see Ryan, in press; Brown, 1980). Another set of problems has to do with subject sampling and experimental design and measurement. Good discussions of these problems have been available for many years (Campbell & Stanley, 1963; R. Thorndike, 1963), but inadequate procedures continue to appear in published studies, and results from these studies continue to be accepted in review articles. Properly sampling subjects, establishing the discriminative power of measures, and avoiding confounding due to statistical regression, require a large commitment of time and resources.

Even when sampling, design, and measurement procedures are adequate, there are serious problems in interpreting the results of studies comparing good and poor readers. We are interested in differences that provide information about the causes of reading disability, but differences obtained in these studies do not necessarily reflect causal factors. Training studies are the best hope for separating causal from noncausal differences.

Finally, there is a critical problem that cannot be remedied by improvements in studies comparing good and poor readers. Interpreting the results of these studies requires the assumption that there is homogeneity within each group of readers; i.e., that certain cognitive components are typically responsible for reading difficulties. However, there may well be a great diversity of patterns of reading disability. Group comparisons would obscure this diversity (Applebee, 1971). The homogeneity assumption can be tested only by detailed studies of individual readers. The lack of studies of individual poor readers, from a cognitive processing point of view of reading, is a critical gap in research. Such studies may make important contributions to our knowledge of reading disability.

References

Applebee, A. N. Research in reading retardation: Two critical problems. Journal of Child Psychology and Psychiatry, 1971, 12, 91-113.

Anderson, R. C. Schema-directed processes in language comprehension (Tech. Rep. No. 50). Urbana-Champaign: University of Illinois, Center for the Study of Reading, July 1977. (ERIC Document Reproduction Service No. ED 142 977)

Barron, R. W. Development of visual word recognition: A review. In T. G. Waller & G. E. MacKinnon (Eds.), Reading research: Advances in theory and practice (Vol. 2). New York: Academic Press, in press.

Bellack, A. A., Davitz, J. R., Kliebard, H. M., & Hyman, R. T. The language of the classroom. New York: Teachers College Press, 1963.

Belmont, J. M., & Butterfield, E. C. The instructional approach to developmental cognitive research. In R. V. Kail & J. W. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, N.J.: Erlbaum, 1977.

Benton, A. L., & Pearl, D. (Eds.), Dyslexia: An appraisal of current knowledge. New York: Oxford University Press, 1978.

Bolinger, D. Accent is predictable (if you're a mind reader). Language, 1972, 48, 633-644.

Bolinger, D. Aspects of language. New York: Harcourt Brace Jovanovich, 1975.

Brown, A. L. Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, N.J.: Erlbaum, 1980.

Brown, A. L., & Campione, J. C. Permissible inferences from the outcome of training studies in cognitive development research. Quarterly Newsletter of the Institute for Comparative Human Development, 1979, 2, 46-53.

Brown, A. L., & Smiley, S. S. Rating the importance of structural units of prose passages: A problem of metacognitive development. Child Development, 1977, 48, 1-8. (a)

Brown, A. L., & Smiley, S. S. The development of strategies for studying prose passages (Tech. Rep. No. 66). Urbana-Champaign: University of Illinois, Center for the Study of Reading, October 1977. (b) (ERIC Document Reproduction Service No. ED 145 371)

Calfee, R. C. Review of Gibson and Levin's 'The Psychology of Reading.' Proceedings of the National Academy of Education, 1976, 3, 1-80.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research in teaching. In N. L. Gage (Ed.), Handbook for research in teaching. Chicago: Rand McNally, 1963. Pp. 171-246.

Carpenter, P. A., & Just, M. A. Reading comprehension as eyes see it. In M. Just & P. Carpenter (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1978.

Chafe, W. L. Language and consciousness. Language, 1974, 50, 111-133.

Chafe, W. Giveness, contrastiveness, definiteness, subjects, topics, and points of view. In C. Li (Ed.), Subject and topic. New York: Academic Press, 1976.

Chapman, L. J., & Chapman, J. P. Problems in the measurement of cognitive deficit. Psychological Bulletin, 1973, 79, 380-385.

Chapman, L. J., & Chapman, J. P. Alternative to the design of manipulating a variable to compare retarded and nonretarded subjects. American Journal of Mental Deficiency, 1974, 79, 404-411.

Clark, H. H., & Clark, E. V. Psychology and language: An introduction to psycholinguistics. New York: Harcourt Brace Jovanovich, 1977.

Clark, H. H., & Haviland, S. E. Comprehension and the given-new contract. In R. O. Freedle (Ed.), Discourse processes and comprehension. Norwood, N.J.: Ablex, 1977.

Clay, M. M., & Imlach, R. H. Juncture, pitch, and stress as reading behavior variables. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 133-139.

Cooper, W. E., & Sorenson, J. M. Fundamental frequency contours and syntactic boundaries. Journal of the Acoustical Society of America, 1977, 62, 133-139.

Cosgrove, J. M., & Patterson, C. J. Plans and the development of listener skills. Developmental Psychology, 1977, 13, 557-564.

Cromer, W. The difference model: A new explanation for some reading difficulties. Journal of Educational Psychology, 1970, 61, 471-483.

Cronbach, L. J., & Snow, C. E. Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington, 1977.

Danks, J. H. Comprehension in listening and reading: Same or different? In Proceedings of the Interdisciplinary Institute on Reading and Child Development. University of Delaware, 1974.

Darwin, C. J. On the dynamic use of prosody in speech perception. In A. Cohen & S. G. Nooteboom (Eds.), Structure and process in speech perception. New York: Springer-Verlag, 1975.

Eammon, D. B. Selection and recall of topical information in prose by better and poorer readers. Reading Research Quarterly, 1978-79, 15, 30-48.

Farr, R. Reading: What can be measured? Newark, Del.: International Reading Association, 1969.

Fleisher, L. S., Jenkins, J. R., & Pany, D. Effects on poor readers' comprehension of training in rapid decoding. Reading Research Quarterly, 1979, 15, 30-48.

Fodor, J. A., & Bever, T. G. The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414-420.

Garnes, S., & Bond, Z. Slips of the ear: Errors in perception of casual speech. In R. E. Grossman, J. J. San, & T. J. Vance (Eds.), Papers from the Eleventh Regional Meeting of the Chicago Linguistics Society. Chicago: Chicago Linguistics Society, 1975.

Geers, A. E. Intonation contour and syntactic structure as predictors of apparent segmentation. Journal of Experimental Psychology: Human Perception and Performance, 1978, 4, 273-283.

Gibson, E. J., & Levin, H. The psychology of reading. Cambridge, Mass.: MIT Press, 1975.

Golinkoff, R. M. A comparison of reading comprehension processes in good and poor reading. Reading Research Quarterly, 1975-76, 4, 623-659.

Grosjean, F., Grosjean, L., & Lane, H. The patterns of silence: Performance structures in sentence production. Cognitive Psychology, 1979, 11, 58-81.

Hornby, P. A. Surface structure and the topic-comment distinction: A developmental study. Child Development, 1971, 42, 975-988.

Hornby, P. A. The psychological subject and predicate. Cognitive Psychology, 1972, 3, 632-642.

Hornby, P. A., & Hass, W. A. Use of contrastive stress by preschool children. Journal of Speech and Hearing Research, 1970, 13, 395-399.

Huey, E. D. The psychology and pedagogy of reading. Cambridge, Mass.: MIT Press, 1968. (Originally published in 1908.)

Ironsmith, M., & Whitehurst, G. J. The development of listener abilities in communication: How children deal with ambiguous information. Child Development, 1978, 49, 348-352.

Jerrolds, B. W., Calloway, B., & Gwaltney, A. A comparative study of three tests of intellectual potential, three tests of reading achievement, and the discrepancy scores between potential and achievement. Journal of Educational Research, 1971, 65, 168-172.

Just, M. A., & Carpenter, P. A. A theory of reading: From eye movements to comprehension. Psychological Review, 1980.

Klatt, D. H. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. Journal of the Acoustical Society of America, 1976, 59, 1208-122.

Kleiman, G. M. Speech recoding in reading. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 323-339.

Kleiman, G. M., Winograd, P. N., & Humphrey, M. M. Prosody and children's parsing of sentences (Tech. Rep. No. 123). Urbana-Champaign: University of Illinois, Center for the Study of Reading, May 1979. (ERIC Document Reproduction Service No. ED 170 733)

Knights, R., & Bakkar, D. J. (Eds.), The neuropsychology of learning disabilities. Baltimore: University Park Press, 1976.

LaBerge, D., & Samuels, S. J. Toward a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 293-323.

Lefevre, C. A. Linguistics, English and the language arts. New York: Teachers College Press, 1970.

Lehiste, I. Phonetic disambiguation of syntactic ambiguity. Glossa, 1973, 7, 107-122.

Levin, J. R. Inducing comprehension in poor readers: A test of recent models. Journal of Educational Psychology, 1973, 65, 10-24.

Lieberman, P. Some acoustic correlates of word stress in American English. Journal of the Acoustical Society of America, 1960, 32, 451-455.

Lieberman, P. Some effects of semantic and grammatical context on the production and perception of speech. Language and Speech, 1963, 6, 172-187.

MacGinitie, W. H. (Ed.), Assessment problems in reading. Newark, Del.: International Reading Association, 1973.

Maratsos, M. The effects of stress on the understanding of pronominal co-reference in children. Journal of Psycholinguistic Research, 1973, 2, 1-8.

Markman, E. M. Realizing that you don't understand: Elementary school children's awareness of inconsistencies. Child Development, 1979.

Meichenbaum, D. Cognitive factors as determinates of learning disabilities: A cognitive functional approach. In R. Knight & D. Bakker (Eds.), The neuropsychology of learning disabilities: Theoretical approaches. Baltimore: University Park Press, 1976.

Neville, M. H., & Pugh, A. K. Context in reading and listening: Variations in approach to cloze tasks. Reading Research Quarterly, 1976-77, 12, 13-31.

Newell, A., & Simon, H. A. Human problem solving. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Olson, D. R. The language of instruction: On the literate bias of schooling. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Erlbaum, 1976.

Olson, D. R. From utterance to text: The bias of language in speech and writing. Harvard Educational Review, 1977, 47, 257-281.

Patterson, C. J., Massad, C. M., & Cosgrove, J. Children's referential communication: Components of plans for effective listening. Developmental Psychology, 1978, 14, 401-406.

Perfetti, C. A., Goldman, S. R., & Hogaboam, T. W. Reading skills and the identification of words in discourse context. Memory & Cognition, 1979, 7, 273-282.

Rayner, K., & McConkie, G. W. What guides a reader's eye movements? Vision Research, 1976, 16, 829-832.

Read, C., Schreiber, P., & Walia, J. Why short subjects are harder to find than long ones. In L. Gleitman & E. Wanner (Eds.), Language acquisition: The state of the art, in press.

Rubin, A. D. Comprehension processes in oral and written language. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, N.J.: Erlbaum, 1980.

Rumelhart, D. E. Toward an interactive theory of reading. In S. Dornic (Ed.), Attention and performance VI. Hillsdale, N.J.: Erlbaum, 1977.

Rutter, M. Prevalence and types of dyslexia. In A. L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge. New York: Oxford University Press, 1978.

Rutter, M., & Yule, W. Specific reading retardation. In L. Mann & D. Sabatino (Eds.), The first review of special education (Vol. 2). Penn.: JSE Publishers, 1973.

Rutter, M., & Yule, W. The concept of specific reading retardation. Journal of Child Psychology and Psychiatry, 1975, 16, 181-197.

Ryan, E. B. Identifying and remediating failures in reading comprehension: Toward an instructional approach for poor comprehenders. In T. G. Waller & G. E. MacKinnon (Eds.), Reading research: Advances in theory and practice (Vol. 2). New York: Academic Press, in press.

Sag, I., & Lieberman, M. The intonational disambiguation of indirect speech acts. In R. E. Grossman, L. J. San, & T. J. Vance (Eds.), Papers from the Eleventh Regional Meeting of the Chicago Linguistic Society. Chicago: Chicago Linguistic Society, 1975.

Samuels, S. J. Success and failure in learning to read: A critique of the research. Reading Research Quarterly, 1973-74, 8, 200-239.

Sartre, J. The words (B. Frechtman, trans.). New York: Braziller, 1964.

Schallert, D. L., & Kleiman, G. M. Some reasons why teachers are easier to understand than textbooks (Reading Education Report No. 9). Urbana-Champaign: University of Illinois, Center for the Study of Reading, July 1979. (ERIC Document Reproduction Service No. ED 174 948)

Scholes, R. J. Acoustic cues for constituent structure. The Hague: Mouton, 1971.

Schvaneveldt, R., Ackerman, B. P., & Semelar, T. The effect of semantic context on children's word recognition. Child Development, 1977, 48, 612-616.

Sinclair, J. M., & Coulthard, R. M. Towards an analysis of discourse: The English used by teachers and pupils. London: Oxford University Press, 1975.

Smith, F. Psycholinguistics and reading. New York: Holt, Rinehart & Winston, 1973.

Snow, C. E., & Ferguson, C. A. Talking to children: Language input and acquisition. Cambridge: Cambridge University Press, 1977.

Sorenson, J. M., Cooper, W. E., & Paccia, J. M. Speech timing of grammatical categories. Cognition, 1978, 6, 135-153.

Spencer, H. The philosophy of style. In Essays: Moral, political and aesthetic. New York: Appleton, 1881. (Originally published in 1852.)

Spiro, R. J. Etiology of reading comprehension style (Tech. Rep. No. 124). Urbana-Champaign: University of Illinois, Center for the Study of Reading, May 1979. (ERIC Document Reproduction Service No. ED 170 734)

Stanovich, K. E. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Reading Research Quarterly, 1980, 16, 32-71.

Taylor, E. A. The spans: Perception, apprehension, and recognition. American Journal of Ophthalmology, 1957, 44, 501-507.

Thorndike, E. L. Reading as reasoning: A study of mistakes in paragraph reading. Journal of Educational Psychology, 1917, 8, 323-332.

Thorndike, R. L. The concepts of over- and underachievement. New York: Teachers College Press, 1963.

Thorndike, R. L., & Hagan, E. Measurement and evaluation in psychology and education. New York: Wiley, 1977.

Tinker, M. A. Recent studies of eye movements in reading. Psychological Bulletin, 1958, 55, 215-23.

Torgesen, J. K. Performance of reading disabled children on serial memory tasks: A selective review of recent research. Reading Research Quarterly, 1978-79, 14, 57-87.

Vellutino, F. R. Dyslexia: Theory and research. Cambridge, Mass.: MIT Press, 1979.

Vygotsky, L. S. Thought and language. Cambridge, Mass.: MIT Press, 1962.

Waller, T. G., & MacKinnon (Eds.), Reading research: Advances in theory and practice (Vol. 2). New York: Academic Press, in press.

Wanat, S. F. Linguistic structure and visual attention in reading. Newark, Del.: International Reading Association, 1971.

Weaver, P. A. Improving reading comprehension: Effects of sentence organization instruction. Reading Research Quarterly, 1979, 15, 129-146.

West, R. F., & Stanovich, K. E. Automatic contextual facilitation in readers of three ages. Child Development, 1978, 49, 717-727.

Wilkinson, A. The foundations of language: Talking and reading in young children. London: Oxford University Press, 1971.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Wingfield, A. The intonation-syntax interaction. In A. Cohen & S. G. Nooteboom (Eds.), Structure and process in speech perception. New York: Springer-Verlag, 1975.

Wingfield, A., & Klein, J. F. Syntactic structure and acoustic pattern in speech perception. Perception & Psychophysics, 1971, 9, 23-25.

Yule, W., Rutter, M., Berger, M., & Thompson, J. Over- and underachievement in reading: Distribution in the general population. British Journal of Educational Psychology, 1974, 44, 1-12.

Footnotes

[1]See Ryan (in press) for a review of the studies that have compared good and poor readers on higher-order linguistic processes and on executive processes.

[2]This statement rests on certain conditions, such as that the measures are equal in discriminative power and that ceiling or floor effects are not responsible for the groups being equal on the baseline task.